Combining neuropsychological and neuroimaging data to assist the early diagnosis of dementia

Submission Number:

633

Authors:

Fermín Segovia[1], Christine Bastin[1], Eric Salmon[1], Christophe Phillips[1]

Institutions:

[1]Cyclotron Research Centre, University of Liège, Liège, Belgium

First Author:

Fermín Segovia - Lecture Information | Contact Me
Cyclotron Research Centre, University of Liège
Liège, Belgium

Introduction:

Alzheimer's disease (AD) is one of the most common neurodegenerative disorders in elderly and it is expected that its prevalence increases in the near future, mainly due to the aging population in developed nations. In recent years, many computer-aided diagnosis (CAD) systems for AD have been presented. Some of these systems analyze neurological brain images by means of machine learning algorithms in order to find the patterns that characterize the disorder, and a few combine several imaging modalities to improve the diagnostic accuracy. However, they usually do not use neuropsychological testing data in that analysis. In this work, we estimated the advantages of combining neuroimaging and neuropsychological testing data in the development of CAD systems based on machine learning.

Methods:

Neuroimages usually contain a huge amount of data compared with the data provided by the neuropsychological testing. In addition, the number of samples used to train the classifier is typically much smaller than the dimension of the neuroimages, which results in the small sample size problem. In order to address these issues, an initial dimensionality reduction step was performed. Principal Component Analysis (PCA), Partial Least Squares (PLS) and Independent Component Analysis (ICA) are three classical techniques used in this work to carry out the dimensionality reduction of the neuroimaging data.
Once the neuroimages were transformed into a reduced set of features, they were combined with neuropsychological testing data using three different approaches:

• Early integration. Combines both data sources before the classification step. Neuroimage features and neuropsychological scores are concatenated resulting in a single feature vector per subject.

• Intermediate integration. Uses a multiple kernel learning approach with a kernel matrix per data source. A linear weighted function is then used to combine the kernel matrices inside the classifier.

• Late integration. Two classifiers are used to individually classify neuroimaging and neuropsychological testing data. The final prediction is estimated by selecting the classifier output with highest confidence.

Classification is performed with a standard linear kernel Support Vector Machine (SVM).

Results:

The database consisted in 46 subjects first diagnosed with Mild Cognitive Impairment (MCI) (Fig. 1). For each subject one FDG-PET scan and five neuropsychological scores were acquired. Three neuropsychological scores were derived from a verbal cued recall memory task, reflecting respectively the efficiency of memory encoding (immediate recall), long-term episodic memory (cued recall) and monitoring capacities (intrusions). The other two neuropsychological scores were phonemic (letter P) and semantic (animals) verbal fluency measures, as an index of executive functioning.
After 3 years of neuropsychological monitored, the diagnosis of some patients was changed to AD. This final diagnosis was used to label the initial data, PET images and neuropsychological scores, as "MCI stable" or "MCI become AD".
All combinations of dimensionality reduction methods and data combination approaches (even using the PET and neuropsychological scores alone) were considered and applied on the labelled data. The accuracy of each system was estimated by means of a cross-validation scheme, Fig. 2. The statistical difference between the accuracy rates obtained when neuropsychological scores are combined, or not, with the imaging data was also assessed by means of a non-parametric test, Fig. 3.

	MCI stable	MCI become AD	Between-group differences
Age	65.55 ± 7.76	72.42 ± 5.91	$t(44) = 3.41$, $p < 0.01$
Education (years)	11.95 ± 3.44	11.38 ± 4.40	$t(44) = -0.47$, $p = 0.63$
MMSE	27.10 ± 1.62	25.26 ± 2.76	$t(44) = -2.63$, $p < 0.05$
Gender (M/F)	10/10	11/15	$\mathrm{Chi}^2 = 0.27$, $p = 0.60$

Figure 1. Demographic details of the subjects that participated in the study used in this work.

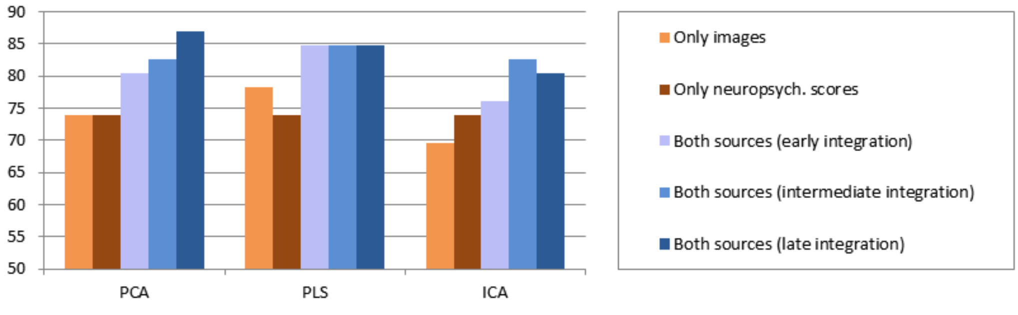

Figure 2. Accuracy rates achieved by the developed CAD systems.

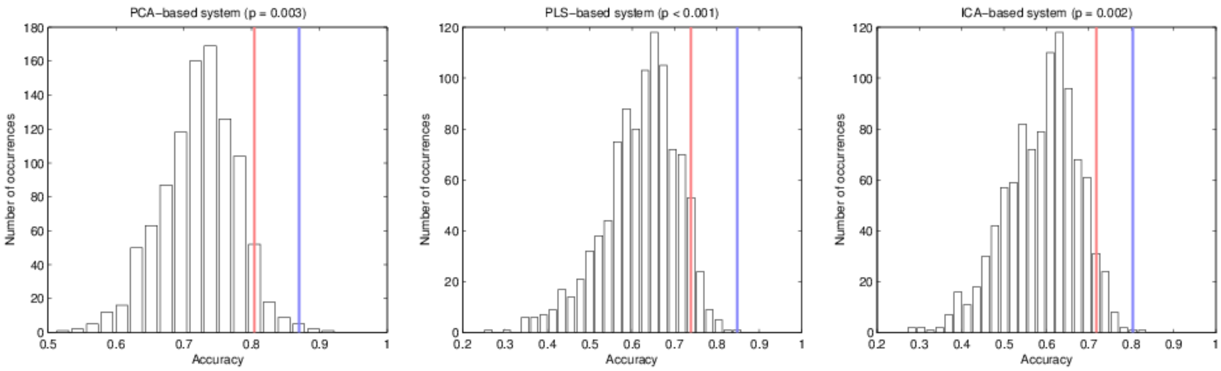

Figure 3. Non-parametric test for the late integration approach: 1000 sets of random neuropsychological scores (same range as the original ones) were generated, then classifier was trained with these random scores (and the image features) and the accuracy estimated. *p*–values were calculated as the (few) number of cases where the accuracy obtained with the random scores was larger than that obtained with the true scores, divided by 1000, i.e. as the probability of obtaining a better accuracy with a random score. Red lines are the accuracies associated with a p-value of 0.05. Blue lines are the accuracies for the late integration approach reported in Figure 2.

Conclusions:

In light of the results shown in previous figures, we can state that including the information from neuropsychological tests improves the accuracy of the analyzed CAD systems. That improvement is achieved by using different ways of combining the data and hardly depends on the processing applied to the neuroimages, i.e. the dimensionality reduction algorithm used.

Modeling and Analysis Methods:

PET Modeling and Analysis

Reference

Brookmeyer, R. et al. (2007), 'Forecasting the global burden of Alzheimer’s disease', Alzheimer’s and Dementia vol. 3, no. 3, pp. 186-191.

Hinrichs, C. et al. (2011), 'Predictive markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population', NeuroImage Vol. 55 no. 2, pp. 574–589.

Illán, I.A. et al. (2011) '18F-FDG PET imaging analysis for computer aided Alzheimer’s diagnosis', Information Sciences Vol. 181, no.4, pp. 903–916.

Jolliffe, I.T. (2002), 'Principal Component Analysis', Springer

McKhann, G.M. (2011), 'The diagnosis of dementia due to Alzheimer’s disease: recommendations from the national institute on aging-Alzheimer’s association workgroups on diagnostic guidelines for Alzheimer’s disease', Alzheimer’s & dementia: the journal of the Alzheimer's Association, Vol. 7, no. 3, pp. 263-269

Segovia, F. et al. (2012), 'Improved Parkinsonism diagnosis using a partial least squares based approach', Medical Physics Vol. 39, no. 7 pp. 4395-4403

Segovia, F. et al. (2014), 'Combining PET images and neuropsychological test data for automatic diagnosis of Alzheimer’s disease', Plos ONE, vol. X no. X, pp. XX-XX.

Varmuza, K. et al. (2009), 'Introduction to Multivariate Statistical Analysis in Chemometrics', CRC Press.